

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.

Application Serial Number: 10/532,053  
Source: PCT  
Date Processed by STIC: 5-5-05

# ***ENTERED***

# CRF Errors Edited by the STIC Systems Branch

Serial Number: 10/S32,053

CRF Edit Date: 5/5/05  
Edited by: TC

\_\_\_ Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

\_\_\_ Corrected the SEQ ID NO. Sequence numbers edited were:

\_\_\_\_\_

\_\_\_ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

\_\_\_\_\_

☒ Deleted: ☒ invalid beginning/end-of-file text ; \_\_\_ page numbers

\_\_\_ Inserted mandatory headings/numeric identifiers, specifically:

\_\_\_\_\_

\_\_\_ Moved responses to same line as heading/numeric identifier, specifically:

\_\_\_\_\_

\_\_\_ Other:

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_



PCT

## RAW SEQUENCE LISTING

DATE: 05/05/2005

PATENT APPLICATION: US/10/532,053

TIME: 14:55:28

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\05052005\J532053.raw

3 <110> APPLICANT: RNA-LINE OY  
 6 <120> TITLE OF INVENTION: Soluble RNA polymerase protein and methods for the use thereof

8 <130> FILE REFERENCE: HY2PCT  
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/532,053  
 C--> 10 <141> CURRENT FILING DATE: 2005-04-21

(pg. 6)

10 <160> NUMBER OF SEQ ID NOS: 15  
 12 <170> SOFTWARE: PatentIn version 3.1  
 14 <210> SEQ ID NO: 1  
 15 <211> LENGTH: 4206  
 16 <212> TYPE: DNA  
 17 <213> ORGANISM: Neurospora crassa  
 19 <400> SEQUENCE: 1

20	atgaacccta	ttactcctag	gaagaggaat	agccccgctc	aggaaatcat	aaaccggctc	60
22	aataacgact	acaacctggg	cctccagtg	gtcgcagaca	caactctcac	ccccccaccg	120
24	cggaggagc	tggccgagag	tgaagaggat	ttcggctgcc	atgacaagat	ctacagagcc	180
26	ctgaactttc	tctactggcg	gaaggatgac	tccttgaacc	aggcagaagc	caactttctc	240
28	atcgaggcca	aagctgcgag	ctcgaactgg	gtgccccaa	cccacgccga	ccctgacacg	300
30	cttccgtgg	ccaaggaacc	tccccgcgcc	gctactgcgc	gccaacaatg	ggcattgcag	360
32	actgtgttgc	tgcaggtgct	taatagggtt	atgccacctc	ccaataacac	accaggtcga	420
34	acgtttggca	gaactctaag	cggcccaagt	ggcctgagcc	gccccaacctc	taccaacacc	480
36	aaacgcaagg	atgagccccg	caatgtcact	ttcgtctgat	cgcacaaaac	ctcgttgact	540
38	cgtcttgcca	caggtcctcc	tattcacggc	gcgccgatac	ccctaaagtt	ccccgatcca	600
40	gtgaataccg	gttccaaaac	accatctctc	gagagtgaga	atctcaatca	gtgcaccaag	660
42	cgggccaagg	gcaagctgtc	tgataatgtt	gcccgtgccg	ccgccccgcc	cgtgcctatt	720
44	gcgagcgctt	tggacaaggt	accgactcga	aggcatgcca	atacgagaga	tcccacggcg	780
46	acaggtcata	gacgagcgga	ccaggtggat	tcccttgata	catctcaagg	cacttcttat	840
48	ggttcgagtg	tcttcagcgc	ttgccgtcac	aatcagagca	ctaccagag	tagttttgag	900
50	gccccctcct	cacagcccag	agagaagcgg	cctgtggatg	ccacggtcct	tgaggctgga	960
52	cacttgattg	agtctcctag	caaaggaaga	acaaccaagt	cccacataga	taaccagccc	1020
54	ctttcatcgt	cctcccagg	tgaaacttcg	ttcagcactt	actatgagtc	gtttccaagt	1080
56	tccggcgcg	agggcgcaat	tcccagagcc	agtcgctcaa	atggactggc	tccgagcgaa	1140
58	gaaagcgctc	gatctcaggt	tcaagttcat	gccccgggtg	ttgcagctcg	gctgagaaat	1200
60	atttggccga	aatttcccaa	atggctacac	gaagctcctc	tcgctgttgc	atgggaagtt	1260
62	accagactct	ttatgcactg	caaagtagac	ttggaagacg	agagcctggg	cctaaagtac	1320
64	gacccttctc	ggtctaccgc	gcgcgatgtc	acagatatct	ggaagactct	ctaccggctt	1380
66	gatgctttcc	gtggtaaac	ctttccagaa	aagccgccca	acgacgtgtt	cgtgaccgca	1440
68	atgacgggca	actttgagag	caaaggtagt	gcccgttggt	tctctgctgt	tctagactac	1500
70	aatccggaca	actcgccctc	tgcgcccctt	taccttggtg	agctgaagcc	gctcatgttc	1560
72	gagcagggtc	gtcgactcac	ccgtcgggtc	ggtcctgata	ggtttttcga	gatacctata	1620
74	ccgtcgccca	cgagcaccag	cccaagtgtg	ccgcccgggt	tcagcaaaaca	accaggtgcg	1680
76	gtcgaagaag	tcatccagtg	gctcacgatg	gggcaacatt	ctctttagtg	ccgccaatgg	1740
78	cgcgctttct	tcgccaagaa	tgcggtatgc	aggaaacctc	tcagggaggt	ccagctccgc	1800
80	gccgaggacc	cgaaacccat	catcaaggag	agagtccact	tctttgccga	gaccggcatt	1860

## RAW SEQUENCE LISTING

DATE: 05/05/2005

PATENT APPLICATION: US/10/532,053

TIME: 14:55:28

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\05052005\J532053.raw

```

82 acgttccgac ctgatgtggt caagacgaga tctgtcgttc cggcagagga acctgtagag 1920
84 caacggaccg agttcaaagt tagtcaaagt ctggactggc tcctgcaact cgacaacaac 1980
86 acttggcagc cgcacctcaa gttgtttctcc cgtatccagc tcggtctgag taagacatat 2040
88 gccattatga cattggagcc tcaccagatc agacaccaca agaccgatct tctttcacct 2100
90 tcaggcactg gcgaagtgat gaatgacggt gtaggcgcga tgtcgcgaag cgtggccaag 2160
92 aggatacgcg atgttctcgg tttgggtgat gtgccctctg ctgtgcaagg gcggtttggt 2220
94 tcggccaagg gaatgtgggt tategacgtt gacgacacag gcgatgagga ttggatcgag 2280
96 acatacccggt ccagcgcgaa gtgggaatgc gacttcgttg ataaacatca acgtaccctc 2340
98 gaagtccgga gcgtggcttc tgaactgaag tcagctggtc tcaacctaca gctgttacct 2400
100 gtcctggaag atagagccag ggacaagggtg aagatgcgcc aggcaatcgg tgaccgtctt 2460
102 atcaacgatt tgcaacgaca gttcagcgag caaaagcatg ctttgaatcg cccagtggaa 2520
104 tttcgccaat ggggtttacga gagttattcc agtcgcgcaa ctcgagtcag ccacggccgt 2580
106 gtgccttttc ttgctgggct acctgacagt caagaggaga cactgaactt cttgatgaac 2640
108 agtgggttcg atcccaagaa gcaaaagtac ttgcaagaca tcgcctggga tcttcaaaag 2700
110 cggaaatgtg acacgttgaa gtccaagctg aacatccgtg tcggtcgatc agcatacatt 2760
112 tacatgattg ccgatttctg ggggtgtgctt gaggaaaatg aggttcattg cggattctcc 2820
114 tcaaagttca gagacgagga ggagtctttt acactcctat cggactgtga tgtcctcgtg 2880
116 gcgcgatccc cagcccattt ccctagtgat atccaacggg ttcgagcagt cttcaagcca 2940
118 gagctccaca gtctcaagga tgtaatcatc ttcttacta aaggagatgt accgcttgct 3000
120 aagaagctat ctggtggaga ctacgacggc gatatggcct gggctcgtgt ggatccggag 3060
122 atcgtcgatg gtttcgtcaa tgcggaaatg cctctggagc ccgacctgtc taggtacct 3120
124 aagaaggaca aaacgacttt caaacaactt atggcctcac acggcacggg ctacgcgcc 3180
126 aaagagcaga ctacatacga tatgatccag aagagcttcc atttcgccct gcagcccaac 3240
128 ttcttgggca tgtgcactaa ctacaaagaa aggtctctgtt acatcaacaa tagtgtgtct 3300
130 aacaagccgg ccatcattct tagttcactg gtgggaaacc tcgtcgatca gagcaagcaa 3360
132 ggtattgtct ttaacgaagc aagctgggct caattgcgta ggggaactgt tggcgggtgca 3420
134 ttgtcccttc ctgacccaat gtacaagagc gacagttggc tcgggcgcgg agagcctacc 3480
136 cacattattg actacctgaa attctccatc gccaggcctg cgattgacaa ggaactggaa 3540
138 gccttccaca atgccatgaa agcggccaag gatacagaag acggcgctca cttttgggat 3600
140 ccgatctctg ctctctacta cacgttcttc aaggagatta gcgacaagtc gcgatcgcc 3660
142 gactgctat tcacgactct gaagaaccgt atcggcgaag tcgagaaaga atatggcagg 3720
144 ttggtcaaaa acaaggagat gagagacagc aaggaccctt accctgtccg cgtcaaccag 3780
146 gtttatgaaa aatggtgcgc catcacgcct gaggcgatgg acaaaccgg agcaaattat 3840
148 gattctaagg tgatcagggt gctggagctg tccttctctg cggaccgtga gatgaatata 3900
150 tgggcattgc tgagggctag cacggctttc aagctgtact accacaagag cccaagttc 3960
152 gtgtggcaga tggcgggcag acagctcgcg tacattaagg cgcagatgac gagcagacc 4020
154 ggtgaaggcg ccccggcgtt gatgaccgcg ttcattgatg cgggcttgat gccggataag 4080
156 aagtttacga agcagtatgt ggccaggctg gaggcgatg gatcgagta ccctgatccg 4140
158 gaggtctatg aagtgtctgg cgatgatgat tttgatggaa ttggtttcac agggaaatggc 4200
160 gattat 4206
163 <210> SEQ ID NO: 2
164 <211> LENGTH: 1402
165 <212> TYPE: PRT
166 <213> ORGANISM: Neurospora crassa
168 <400> SEQUENCE: 2
170 Met Asn Pro Ile Thr Pro Arg Lys Arg Asn Ser Pro Val Glu Glu Ile
171 1 5 10 15
174 Ile Asn Arg Leu Asn Asn Asp Tyr Asn Leu Gly Leu Gln Cys Val Ala
175 20 25 30

```

## RAW SEQUENCE LISTING

DATE: 05/05/2005

PATENT APPLICATION: US/10/532,053

TIME: 14:55:28

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\05052005\J532053.raw

```

178 Asp Thr Thr Leu Thr Pro His Arg Arg Lys Glu Leu Ala Glu Ser Asp
179          35                      40                      45
182 Glu Asp Phe Gly Arg His Asp Lys Ile Tyr Arg Ala Leu Asn Phe Leu
183          50                      55                      60
186 Tyr Trp Arg Lys Asp Asp Ser Leu Asn Gln Ala Glu Ala Asn Phe Phe
187 65                      70                      75                      80
190 Ile Glu Ala Lys Ala Ala Ser Ser Asn Trp Val Pro Lys Ala His Ala
191          85                      90                      95
194 Asp Pro Asp Thr Leu Pro Trp Ser Lys Glu Pro Pro Arg Ala Ala Thr
195          100                     105                     110
198 Ala Gly Gln Gln Trp Ala Leu Gln Thr Val Leu Leu Glu Val Leu Asn
199          115                     120                     125
202 Arg Phe Met Pro Pro Pro Asn Asn Thr Pro Gly Arg Thr Phe Gly Arg
203          130                     135                     140
206 Thr Leu Ser Gly Pro Ser Gly Leu Ser Arg Pro Thr Ser Thr Asn Thr
207 145                     150                     155                     160
210 Lys Arg Lys Asp Glu Pro Ala Asn Val Thr Phe Ala Asp Pro Pro Lys
211          165                     170                     175
214 Arg Ser Leu Thr Arg Ser Ala Thr Gly Pro Pro Ile His Gly Ala Ala
215          180                     185                     190
218 Ile Pro Leu Lys Phe Pro Asp Pro Val Asn Thr Gly Ser Lys Arg Pro
219          195                     200                     205
222 Ser Leu Glu Ser Glu Asn Leu Asn Gln Cys Thr Lys Arg Ala Lys Gly
223          210                     215                     220
226 Lys Leu Ser Asp Asn Val Ala Ala Ala Ala Ala Pro Pro Val Pro Ile
227 225                     230                     235                     240
230 Ala Ser Ala Leu Asp Lys Val Pro Thr Arg Arg His Ala Asn Thr Arg
231          245                     250                     255
234 Asp Pro Thr Ala Thr Gly His Arg Arg Ala Asp Gln Val Asp Ser Phe
235          260                     265                     270
238 Asp Thr Ser Gln Gly Thr Ser Tyr Gly Ser Ser Val Phe Ser Ala Cys
239          275                     280                     285
242 Arg His Asn Gln Ser Thr Thr Gln Ser Ser Phe Glu Ala Pro Pro Ser
243          290                     295                     300
246 Gln Pro Arg Glu Lys Arg Pro Val Asp Ala Thr Val Phe Glu Ala Gly
247 305                     310                     315                     320
250 His Leu Ile Glu Ser Pro Ser Lys Gly Arg Thr Thr Lys Ser His Ile
251          325                     330                     335
254 Asp Asn Gln Pro Leu Ser Ser Ser Ser Gln Gly Glu Thr Ser Phe Ser
255          340                     345                     350
258 Thr Tyr Tyr Glu Ser Phe Pro Ser Ser Gly Gly Glu Gly Ala Ile Pro
259          355                     360                     365
262 Glu Pro Ser Arg Ser Asn Gly Leu Ala Arg Ser Glu Glu Ser Ala Arg
263          370                     375                     380
266 Ser Gln Val Gln Val His Ala Pro Val Val Ala Ala Arg Leu Arg Asn
267 385                     390                     395                     400
270 Ile Trp Pro Lys Phe Pro Lys Trp Leu His Glu Ala Pro Leu Ala Val
271          405                     410                     415
274 Ala Trp Glu Val Thr Arg Leu Phe Met His Cys Lys Val Asp Leu Glu

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/532,053

DATE: 05/05/2005

TIME: 14:55:28

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\05052005\J532053.raw

```

275          420          425          430
278 Asp Glu Ser Leu Gly Leu Lys Tyr Asp Pro Ser Trp Ser Thr Ala Arg
279          435          440          445
282 Asp Val Thr Asp Ile Trp Lys Thr Leu Tyr Arg Leu Asp Ala Phe Arg
283          450          455          460
286 Gly Lys Pro Phe Pro Glu Lys Pro Pro Asn Asp Val Phe Val Thr Ala
287 465          470          475          480
290 Met Thr Gly Asn Phe Glu Ser Lys Gly Ser Ala Val Val Leu Ser Ala
291          485          490          495
294 Val Leu Asp Tyr Asn Pro Asp Asn Ser Pro Thr Ala Pro Leu Tyr Leu
295          500          505          510
298 Val Lys Leu Lys Pro Leu Met Phe Glu Gln Gly Cys Arg Leu Thr Arg
299          515          520          525
302 Arg Phe Gly Pro Asp Arg Phe Phe Glu Ile Leu Ile Pro Ser Pro Thr
303          530          535          540
306 Ser Thr Ser Pro Ser Val Pro Pro Val Val Ser Lys Gln Pro Gly Ala
307 545          550          555          560
310 Val Glu Glu Val Ile Gln Trp Leu Thr Met Gly Gln His Ser Leu Val
311          565          570          575
314 Gly Arg Gln Trp Arg Ala Phe Phe Ala Lys Asp Ala Gly Tyr Arg Lys
315          580          585          590
318 Pro Leu Arg Glu Phe Gln Leu Arg Ala Glu Asp Pro Lys Pro Ile Ile
319          595          600          605
322 Lys Glu Arg Val His Phe Phe Ala Glu Thr Gly Ile Thr Phe Arg Pro
323          610          615          620
326 Asp Val Phe Lys Thr Arg Ser Val Val Pro Ala Glu Glu Pro Val Glu
327 625          630          635          640
330 Gln Arg Thr Glu Phe Lys Val Ser Gln Met Leu Asp Trp Leu Leu Gln
331          645          650          655
334 Leu Asp Asn Asn Thr Trp Gln Pro His Leu Lys Leu Phe Ser Arg Ile
335          660          665          670
338 Gln Leu Gly Leu Ser Lys Thr Tyr Ala Ile Met Thr Leu Glu Pro His
339          675          680          685
342 Gln Ile Arg His His Lys Thr Asp Leu Leu Ser Pro Ser Gly Thr Gly
343          690          695          700
346 Glu Val Met Asn Asp Gly Val Gly Arg Met Ser Arg Ser Val Ala Lys
347 705          710          715          720
350 Arg Ile Arg Asp Val Leu Gly Leu Gly Asp Val Pro Ser Ala Val Gln
351          725          730          735
354 Gly Arg Phe Gly Ser Ala Lys Gly Met Trp Val Ile Asp Val Asp Asp
355          740          745          750
358 Thr Gly Asp Glu Asp Trp Ile Glu Thr Tyr Pro Ser Gln Arg Lys Trp
359          755          760          765
362 Glu Cys Asp Phe Val Asp Lys His Gln Arg Thr Leu Glu Val Arg Ser
363          770          775          780
366 Val Ala Ser Glu Leu Lys Ser Ala Gly Leu Asn Leu Gln Leu Leu Pro
367 785          790          795          800
370 Val Leu Glu Asp Arg Ala Arg Asp Lys Val Lys Met Arg Gln Ala Ile
371          805          810          815

```

## RAW SEQUENCE LISTING

DATE: 05/05/2005

PATENT APPLICATION: US/10/532,053

TIME: 14:55:28

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\05052005\J532053.raw

```

374 Gly Asp Arg Leu Ile Asn Asp Leu Gln Arg Gln Phe Ser Glu Gln Lys
375           820           825           830
378 His Ala Leu Asn Arg Pro Val Glu Phe Arg Gln Trp Val Tyr Glu Ser
379           835           840           845
382 Tyr Ser Ser Arg Ala Thr Arg Val Ser His Gly Arg Val Pro Phe Leu
383           850           855           860
386 Ala Gly Leu Pro Asp Ser Gln Glu Glu Thr Leu Asn Phe Leu Met Asn
387 865           870           875           880
390 Ser Gly Phe Asp Pro Lys Lys Gln Lys Tyr Leu Gln Asp Ile Ala Trp
391           885           890           895
394 Asp Leu Gln Lys Arg Lys Cys Asp Thr Leu Lys Ser Lys Leu Asn Ile
395           900           905           910
398 Arg Val Gly Arg Ser Ala Tyr Ile Tyr Met Ile Ala Asp Phe Trp Gly
399           915           920           925
402 Val Leu Glu Glu Asn Glu Val His Val Gly Phe Ser Ser Lys Phe Arg
403           930           935           940
406 Asp Glu Glu Glu Ser Phe Thr Leu Leu Ser Asp Cys Asp Val Leu Val
407 945           950           955           960
410 Ala Arg Ser Pro Ala His Phe Pro Ser Asp Ile Gln Arg Val Arg Ala
411           965           970           975
414 Val Phe Lys Pro Glu Leu His Ser Leu Lys Asp Val Ile Ile Phe Ser
415           980           985           990
418 Thr Lys Gly Asp Val Pro Leu Ala Lys Lys Leu Ser Gly Gly Asp Tyr
419           995           1000           1005
422 Asp Gly Asp Met Ala Trp Val Cys Trp Asp Pro Glu Ile Val Asp
423           1010           1015           1020
426 Gly Phe Val Asn Ala Glu Met Pro Leu Glu Pro Asp Leu Ser Arg
427           1025           1030           1035
430 Tyr Leu Lys Lys Asp Lys Thr Thr Phe Lys Gln Leu Met Ala Ser
431           1040           1045           1050
434 His Gly Thr Gly Ser Ala Ala Lys Glu Gln Thr Thr Tyr Asp Met
435           1055           1060           1065
438 Ile Gln Lys Ser Phe His Phe Ala Leu Gln Pro Asn Phe Leu Gly
439           1070           1075           1080
442 Met Cys Thr Asn Tyr Lys Glu Arg Leu Cys Tyr Ile Asn Asn Ser
443           1085           1090           1095
446 Val Ser Asn Lys Pro Ala Ile Ile Leu Ser Ser Leu Val Gly Asn
447           1100           1105           1110
450 Leu Val Asp Gln Ser Lys Gln Gly Ile Val Phe Asn Glu Ala Ser
451           1115           1120           1125
454 Trp Ala Gln Leu Arg Arg Glu Leu Leu Gly Gly Ala Leu Ser Leu
455           1130           1135           1140
458 Pro Asp Pro Met Tyr Lys Ser Asp Ser Trp Leu Gly Arg Gly Glu
459           1145           1150           1155
462 Pro Thr His Ile Ile Asp Tyr Leu Lys Phe Ser Ile Ala Arg Pro
463           1160           1165           1170
466 Ala Ile Asp Lys Glu Leu Glu Ala Phe His Asn Ala Met Lys Ala
467           1175           1180           1185
470 Ala Lys Asp Thr Glu Asp Gly Ala His Phe Trp Asp Pro Asp Leu

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/532,053

DATE: 05/05/2005  
TIME: 14:55:29

Input Set : A:\pto.kd.txt  
Output Set: N:\CRF4\05052005\J532053.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:13; Xaa Pos. 8

**VERIFICATION SUMMARY**

DATE: 05/05/2005

PATENT APPLICATION: US/10/532,053

TIME: 14:55:29

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\05052005\J532053.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:1017 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0